ANASIEZE IKENNA – CLOUD & AI ENGINEER

Project: Real-Time Streaming Analytics Into BigQuery

Overview

This project demonstrates how I designed and deployed a real time streaming analytics pipeline on Google Cloud. The system ingests live environmental sensor data, processes it in motion, and stores it in BigQuery for immediate querying and analysis.

Rather than relying on batch jobs or manual uploads, the architecture supports continuous data flow, making it suitable for IoT sensors, monitoring systems, and real time analytics workloads.

Problem Statement

A newly formed development team needed a way to:

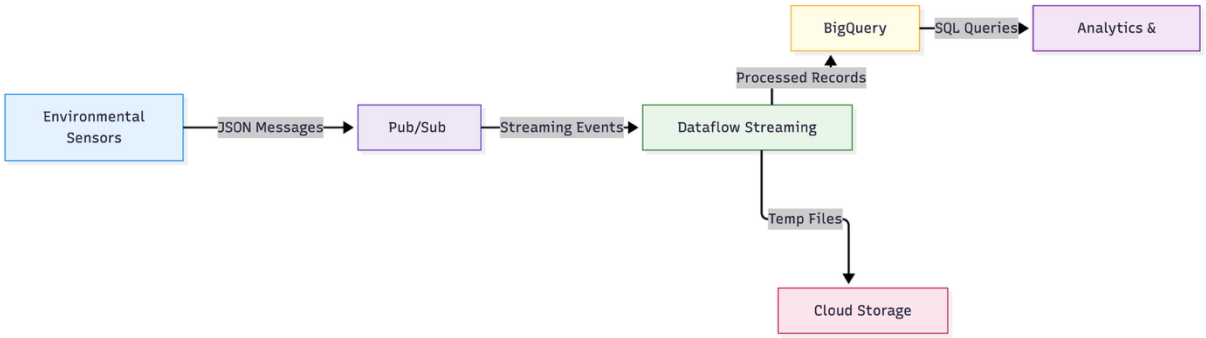
- Ingest real time temperature data from sensors
- Process streaming events reliably at scale
- Store the data in an analytics ready warehouse
- Validate incoming data instantly

Goal: I was responsible for designing and implementing the end to end streaming pipeline using Google Cloud managed services.

Tech Stack

- Cloud Storage for Dataflow temporary and staging storage
- Pub/Sub for event ingestion
- Dataflow for stream processing
- BigQuery for analytics

Architecture Overview



Flow of Data

1. Environmental sensor data is published as events
2. Pub/Sub handles real time message ingestion
3. Dataflow processes the stream continuously
4. BigQuery stores the streaming data for analytics
5. SQL queries validate and analyze incoming records

This architecture is serverless, scalable, and production aligned.

Deployment Steps

+ **Provisioned Cloud Storage for Dataflow execution job**

+ **Created a BigQuery dataset and table for streaming ingestion**

+ **Configured Pub/Sub for real time message ingestion**

+ **Deployed a Dataflow streaming pipeline from Pub/Sub to BigQuery**

+ **Published test events and validated live ingestion using BigQuery queries**

Each component was created in the correct region with required APIs enabled to reflect production best practices.

Outcome

- A fully running real time streaming pipeline
- Live data flowing from Pub/Sub into BigQuery
- A reproducible, cloud native analytics architecture

This project shows I can design, deploy, and operate real time cloud data pipelines, troubleshoot distributed systems independently, and think end to end like a cloud engineer.